SVM Soft Margin Classification

Nipun Batra

July 20, 2025

IIT Gandhinagar

null

Soft-Margin SVM

- Can we learn SVM for “slightly” non-separable data without projecting to a higher space?

Soft-Margin SVM

- Can we learn SVM for “slightly” non-separable data without projecting to a higher space?
- Introduce some “slack” (ξ_i) or loss or penalty for samples - allow some samples to be misclassified

null

Soft-Margin SVM

Change Objective

$\min \frac{1}{2}||\bar{w}||^2 + C \sum_{i=1}^{n} \xi_i$

s.t. $y_i(\bar{w}\bar{x}_i + b) \geq 1 - \xi_i$

Soft-Margin SVM

Change Objective

$\min \frac{1}{2}||\bar{w}||^2 + C\sum_{i=1}^{n}\xi_i$
s.t. $y_i(\bar{w}\bar{x}_i + b) \geq 1 - \xi_i$

In Dual:

$$\text{Minimize} \sum_{i=1}^{n}\alpha_i - \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j \bar{x}_i \bar{x}_j$$

s.t.

$$0 \leq \alpha_i \leq C \quad \& \quad \sum_{i=1}^{n}\alpha_i y_i = 0$$

null

Bias Variance Trade-off for Soft-Margin SVM

Low C $\implies$ Higher train error (higher bias)

High C $\implies$ Very sensitive to datasete (high variance)

Soft-Margin SVM

If C $\rightarrow 0$

Objective $\rightarrow \min \frac{1}{2}||\bar{w}||^2$

$\implies$ Choose large margin (without worrying for ξ_is)

$$\boxed{\text{Recall: Margin} = \frac{2}{||\bar{w}||}}$$

If C $\rightarrow \infty$ (or very large) Objective $\rightarrow \min C \sum \xi_i$ or choose W, b, s.t. ξ_i is small!

Soft-Margin SVM

Q) What is the equivalent of hard margin?

a $C \to 0$

Soft-Margin SVM

Q) What is the equivalent of hard margin?

a $C \to 0$

b $C \to \infty$

Soft-Margin SVM

Q) What is the equivalent of hard margin?

a $C \to 0$

Soft-Margin SVM

Q) What is the equivalent of hard margin?

a $C \to 0$

b $\boxed{C \to \infty} \implies$ No violations!!

Soft-Margin SVM

Types of support vectors:

- Zone 2: $y_i(\bar{w}\bar{x}_i + b) = 1$

$\therefore$ As C increases, # support vectors decreases

Notebook: SVM-soft-margin

Soft-Margin SVM

Types of support vectors:

- Zone 2: $y_i(\bar{w}\bar{x}_i + b) = 1$
- Zone 3: $0 < \xi_i < 1$ (correctly classified)

$\therefore$ As C increases, $\#$ support vectors decreases

Notebook: SVM-soft-margin

Soft-Margin SVM

Types of support vectors:

- Zone 2: $y_i(\bar{w}\bar{x}_i + b) = 1$
- Zone 3: $0 < \xi_i < 1$ (correctly classified)
- Zone 4: $\xi_i > 1$ (Misclassified)

$\therefore$ As C increases, $\#$ support vectors decreases

Notebook: SVM-soft-margin

SVM Formulation in the Loss + Penalty Form

Objective:

$$\min \frac{1}{2}||\bar{w}||^2 + C\sum_{i=1}^{N}\xi_i$$

Now:

$$y_i(\bar{w}\bar{x}_i + b) \geq 1 - \xi_i$$

$$\xi_i \geq 1 - y_i(\bar{w}\bar{x}_i + b)$$

But $\xi_i \geq 0$

$$\therefore \xi_i = \max\left[0, 1 - y_i(\bar{w}\bar{x}_i + b)\right]$$

SVM Formulation in the Loss + Penalty Form

$\therefore$ Objective is:

$$\min C \sum \xi_i + \frac{1}{2}||\bar{w}||^2$$

$$\implies \min C \sum_{i=1}^{N} \max\left[0, 1 - y_i(\bar{w}\bar{x}_i + b)\right] + \frac{1}{2}||\bar{w}||^2$$

$$\implies \min \underbrace{\sum_{i=1}^{N} \max\left[0, 1 - y_i(\bar{w}\bar{x}_i + b)\right]}_{\text{Loss}} + \underbrace{\frac{1}{2C}}_{\text{Regularisation}}||\bar{w}||^2$$

null

Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\bar{w}\bar{x}_i + b)\right]$

- Case I $\quad y_i(\bar{w}\bar{x}_i + b) = 1$
  Lies on Margin: $Loss_i = 0$

Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\bar{w}\bar{x}_i + b)\right]$

- Case I $\qquad y_i(\bar{w}\bar{x}_i + b) = 1$
  Lies on Margin: $Loss_i = 0$
- Case II
  $$y_i(\bar{w}\bar{x}_i + b) > 1$$
  $Loss_i = 0$

Loss Function for Sum (Hinge Loss)

Loss function is $\sum_{i=1}^{N} \max\left[0, 1 - y_i(\bar{w}\bar{x}_i + b)\right]$

- Case I $\quad y_i(\bar{w}\bar{x}_i + b) = 1$
  Lies on Margin: $Loss_i = 0$
- Case II
  $$y_i(\bar{w}\bar{x}_i + b) > 1$$
  $Loss_i = 0$
- Case III
  $$y_i(\bar{w}\bar{x}_i + b) < 1$$
  $Loss_i \neq 0$

Hinge Loss Continued

Q) Is hinge loss convex and differentiable?

Convex: ✓

Differentiable: X

Subgradient: ✓

SVM Loss is Convex

Hinge Loss $\sum(\max[0, (1 - y_i(\bar{w}x_i + b))]$ is convex

Penalty $\frac{1}{2}||\bar{w}||^2$ is convex

$\therefore$ SVM loss is convex